



## RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/889,344A  
Source: PCT 09  
Date Processed by STIC: 2/13/02

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.

PATENTIN 2.1 e-mail help: [patin21help@uspto.gov](mailto:patin21help@uspto.gov) or phone 703-306-4119 (R. Wax)

PATENTIN 3.0 e-mail help: [patin3help@uspto.gov](mailto:patin3help@uspto.gov) or phone 703-306-4119 (R. Wax)

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 3.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:

<http://www.uspto.gov/web/offices/pac/checker>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. Hand Carry directly to:  
U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7<sup>th</sup> Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202  
Or  
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

## Raw Sequence Listing Error Summary

<u>ERROR DETECTED</u>	<u>SUGGESTED CORRECTION</u>	<u>SERIAL NUMBER:</u> <i>09/889, 344 A</i>
<b>ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE</b>		
1 <input type="checkbox"/> Wrapped Nucleic Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 <input type="checkbox"/> Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.	
3 <input type="checkbox"/> Misaligned Amino Numbering	The numbering under each 5 <sup>th</sup> amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.	
4 <input type="checkbox"/> Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.	
5 <input type="checkbox"/> Variable Length	Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 <input type="checkbox"/> PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.	
7 <input type="checkbox"/> Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped	
	Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.	
8 <input type="checkbox"/> Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000	
9 <input type="checkbox"/> Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
10 <input type="checkbox"/> Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence	
11 <input checked="" type="checkbox"/> Use of <220>	Sequence(s) <i>1, 2, 3</i> missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 <input type="checkbox"/> PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	



PCT09

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/889,344A

DATE: 02/13/2002

TIME: 18:59:17

Input Set : A:\USSEQLIST2.txt

Output Set: N:\CRF3\02132002\I889344A.raw

Does Not Comply  
Corrected Diskette Needed

```

4 <110> APPLICANT: CHEN, WENFANG
5     MEEK, THOMAS D.
6     POWELL, DAVID J.
7     TEW, DAVID G.
10 <120> TITLE OF INVENTION: Method of Site Specific Labeling of Proteins and Uses
11 Therefor
14 <130> FILE REFERENCE: P50892
16 <140> CURRENT APPLICATION NUMBER: 09/889,344A
17 <141> CURRENT FILING DATE: 2001-07-16
19 <150> PRIOR APPLICATION NUMBER: PCT/US00/01481
20 <151> PRIOR FILING DATE: 2000-01-20
22 <150> PRIOR APPLICATION NUMBER: US 60/117,327
23 <151> PRIOR FILING DATE: 1999-01-22
25 <160> NUMBER OF SEQ ID NOS: 16
27 <170> SOFTWARE: FastSEQ for Windows Version 3.0
29 <210> SEQ ID NO: 1
30 <211> LENGTH: 5
31 <212> TYPE: PRT
32 <213> ORGANISM: Artificial Sequence
34 <220> FEATURE:
35 <221> NAME/KEY: unsure
36 <222> LOCATION: (5)
37 <223> OTHER INFORMATION: Where Xaa at position (5) can represent Leucine or Isoleucine
39 <400> SEQUENCE: 1
W--> 40 Gln Ser Lys Val Xaa
41      1           5
43 <210> SEQ ID NO: 2
44 <211> LENGTH: 207
45 <212> TYPE: PRT
46 <213> ORGANISM: Artificial Sequence
48 <220> FEATURE:
49 <221> NAME/KEY: unsure
50 <222> LOCATION: (1)(2)(3)(4)(5)(6)(7)(8)(9)(10)(11)(12)(13)(14)(15)
51 <222> LOCATION: (16)(17)(18)(19)(20)(21)(22)(23)(24)(25)(26)(27)(28)
52 <222> LOCATION: (29)(30)(31)(32)(33)(34)(35)(36)(37)(38)(39)(40)(41)
53 <222> LOCATION: (42)(43)(44)(45)(46)(47)(48)(49)(50)(51)(52)(53)(54)
54 <222> LOCATION: (55)(56)(57)(58)(59)(60)(61)(62)(62)(64)(65)(66)(67)
55 <222> LOCATION: (68)(69)(70)(71)(72)(73)(74)(75)(76)(77)(78)(79)(80)
56 <222> LOCATION: (81)(82)(83)(84)(85)(86)(87)(88)(89)(90)(91)(92)(93)
57 <222> LOCATION: (94)(95)(96)(97)(98)(99)(100)(101)(106)(107)(108)(109)
58 <222> LOCATION: (110)(111)(112)(113)(114)(115)(116)(117)(118)(119)(120)
59 <222> LOCATION: (121)(122)(123)(124)(125)(126)(127)(128)(129)(130)(131)
60 <222> LOCATION: (132)(133)(134)(135)(136)(137)(138)(139)(140)(141)(142)

    USE of Artificial Sequence must
    be accompanied by feature <220> and
    <223> to explain origin of genetic
    material. See item # 11 on ERROR
    SUMMARY SHEET
```

Note! If all Xaa's  
are equal, then  
you can use a  
range, i.e. (1)...(207),  
for location.

RAW SEQUENCE LISTING DATE: 02/13/2002  
 PATENT APPLICATION: US/09/889,344A TIME: 18:59:17

Input Set : A:\USSEQLIST2.txt  
 Output Set: N:\CRF3\02132002\I889344A.raw

61 <222> LOCATION: (143)(144)(145)(146)(147)(148)(149)(150)(151)(152)(153)  
 62 <222> LOCATION: (154)(155)(156)(157)(158)(159)(160)(161)(162)(163)(164)  
 63 <222> LOCATION: (165)(166)(167)(168)(169)(170)(171)(172)(173)(174)(175)  
 64 <222> LOCATION: (176)(177)(178)(179)(180)(181)(182)(183)(184)(185)(186)  
 65 <222> LOCATION: (187)(188)(189)(190)(191)(192)(193)(194)(195)(196)(197)  
 66 <222> LOCATION: (198)(199)(200)(201)(202)(203)(204)(205)(206)(207)  
 67 <223> OTHER INFORMATION: Where Xaa can represent none or any one of the twenty naturally

naturally  
 68 <223> OTHER INFORMATION: occurring amino acids  
 70 <400> SEQUENCE: 2

W--> 71 Xaa  
 72 1 5 10 15  
 W--> 73 Xaa  
 74 20 25 30  
 W--> 75 Xaa  
 76 35 40 45  
 W--> 77 Xaa  
 78 50 55 60  
 W--> 79 Xaa  
 80 65 70 75 80  
 W--> 81 Xaa  
 82 85 90 95  
 W--> 83 Xaa Xaa Xaa Xaa Gln Ser Lys Val Xaa Xaa Xaa Xaa Xaa Xaa  
 84 100 105 110  
 W--> 85 Xaa  
 86 115 120 125  
 W--> 87 Xaa  
 88 130 135 140  
 W--> 89 Xaa  
 90 145 150 155 160  
 W--> 91 Xaa  
 92 165 170 175  
 W--> 93 Xaa  
 94 180 185 190  
 W--> 95 Xaa  
 96 195 200 205

98 <210> SEQ ID NO: 3

99 <211> LENGTH: 207

100 <212> TYPE: PRT

101 <213> ORGANISM: Artificial Sequence

*- see page 1*

103 <220> FEATURE:

104 <221> NAME/KEY: unsure

105 <222> LOCATION: (1)(2)(3)(4)(5)(6)(7)(8)(9)(10)(11)(12)(13)(14)(15)

106 <222> LOCATION: (16)(17)(18)(19)(20)(21)(22)(23)(24)(25)(26)(27)(28)

107 <222> LOCATION: (29)(30)(31)(32)(33)(34)(35)(36)(37)(38)(39)(40)(41)

108 <222> LOCATION: (42)(43)(44)(45)(46)(47)(48)(49)(50)(51)(52)(53)(54)

109 <222> LOCATION: (55)(56)(57)(58)(59)(60)(61)(62)(63)(64)(65)(66)(67)

110 <222> LOCATION: (68)(69)(70)(71)(72)(73)(74)(75)(76)(77)(78)(79)(80)

111 <222> LOCATION: (81)(82)(83)(84)(85)(86)(87)(88)(89)(90)(91)(92)(93)

112 <222> LOCATION: (94)(95)(96)(97)(98)(99)(100)(101)(106)(107)(108)(109)

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/889,344A

DATE: 02/13/2002  
TIME: 16:59:17

Input Set : A:\USSEQLIST2.txt  
Output Set: N:\CRF3\02132002\I889344A.raw

113 <222> LOCATION: (110)(111)(112)(113)(114)(115)(116)(117)(118)(119)(120)  
114 <222> LOCATION: (121)(122)(123)(124)(125)(126)(127)(128)(129)(130)(131)  
115 <222> LOCATION: (132)(133)(134)(135)(136)(137)(138)(139)(140)(141)(142)  
116 <222> LOCATION: (143)(144)(145)(146)(147)(148)(149)(150)(151)(152)(153)  
117 <222> LOCATION: (154)(155)(156)(157)(158)(159)(160)(161)(162)(163)(164)  
118 <222> LOCATION: (165)(166)(167)(168)(169)(170)(171)(172)(173)(174)(175)  
119 <222> LOCATION: (176)(177)(178)(179)(180)(181)(182)(183)(184)(185)(186)  
120 <222> LOCATION: (187)(188)(189)(190)(191)(192)(193)(194)(195)(196)(197)  
121 <222> LOCATION: (198)(199)(200)(201)(202)(203)(204)(205)(206)(207)

122 <223> OTHER INFORMATION: Where Xaa can represent none or any one of the twenty naturally occurring amino acids

123 <223> OTHER INFORMATION: occurring amino acids  
125 <400> SEQUENCE: 3

W--> 126 Xaa  
127 1 5 10 15  
W--> 128 Xaa  
129 20 25 30  
W--> 130 Xaa  
131 35 40 45  
W--> 132 Xaa  
133 50 55 60  
W--> 134 Xaa  
135 65 70 75 80  
W--> 136 Xaa  
137 85 90 95  
W--> 138 Xaa Xaa Xaa Xaa Xaa Gln Ser Lys Val Xaa Xaa Xaa Xaa Xaa Xaa  
139 100 105 110  
W--> 140 Xaa  
141 115 120 125  
W--> 142 Xaa  
143 130 135 140  
W--> 144 Xaa  
145 145 150 155 160  
W--> 146 Xaa  
147 165 170 175  
W--> 148 Xaa  
149 180 185 190  
W--> 150 Xaa  
151 195 200 205

153 <210> SEQ ID NO: 4

154 <211> LENGTH: 10

155 <212> TYPE: PRT

156 <213> ORGANISM: Artificial Sequence

158 <220> FEATURE:

159 <223> OTHER INFORMATION: Derivative of a factor XIII substrate

161 <400> SEQUENCE: 4

162 Leu Ser Leu Ser Gln Ser Lys Val Leu Gly

163 1 5 10

165 <210> SEQ ID NO: 5

166 <211> LENGTH: 10

RAW SEQUENCE LISTING DATE: 02/13/2002  
PATENT APPLICATION: US/09/889,344A TIME: 18:59:17

Input Set : A:\USSEQLIST2.txt  
Output Set: N:\CRF3\02132002\I889344A.raw

167 <212> TYPE: PRT  
168 <213> ORGANISM: Artificial Sequence  
170 <220> FEATURE:  
171 <223> OTHER INFORMATION: Derivative of a factor XIII substrate  
173 <400> SEQUENCE: 5  
174 Ile Gly Glu Gly Gln Ser Lys Val Leu Gly  
175 1 5 10  
177 <210> SEQ ID NO: 6  
178 <211> LENGTH: 10  
179 <212> TYPE: PRT  
180 <213> ORGANISM: Artificial Sequence  
182 <220> FEATURE:  
183 <223> OTHER INFORMATION: Derivative of a factor XIII substrate  
185 <400> SEQUENCE: 6  
186 Leu Gly Pro Gly Gln Ser Lys Val Ile Gly  
187 1 5 10  
189 <210> SEQ ID NO: 7  
190 <211> LENGTH: 81  
191 <212> TYPE: DNA  
192 <213> ORGANISM: Unknown  
194 <220> FEATURE:  
195 <223> OTHER INFORMATION: Oligonucleotide designed to introduce Q tag  
197 <400> SEQUENCE: 7  
198 tgtacctcag accatatgag cctgtccctg tcccaagtcca aagttctgcc gggtccgagc 60  
199 actatcgaag aacgcgttaa g 81  
201 <210> SEQ ID NO: 8  
202 <211> LENGTH: 37  
203 <212> TYPE: DNA  
204 <213> ORGANISM: Unknown  
206 <220> FEATURE:  
207 <223> OTHER INFORMATION: Oligonucleotide designed to introduce Q tag  
209 <400> SEQUENCE: 8  
210 tgatgtcagt caagcttacg cctggtgcc gttgatg 37  
212 <210> SEQ ID NO: 9  
213 <211> LENGTH: 14  
214 <212> TYPE: PRT  
215 <213> ORGANISM: Artificial Sequence  
217 <220> FEATURE:  
218 <223> OTHER INFORMATION: Derivative of a factor XIII substrate  
220 <400> SEQUENCE: 9  
221 Met Ser Leu Ser Leu Ser Gln Ser Lys Val Leu Pro Gly Pro  
222 1 5 10  
224 <210> SEQ ID NO: 10  
225 <211> LENGTH: 37  
226 <212> TYPE: DNA  
227 <213> ORGANISM: Unknown  
229 <220> FEATURE:  
230 <223> OTHER INFORMATION: Oligonucleotide designed to introduce Q tag  
232 <400> SEQUENCE: 10

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/889,344A

DATE: 02/13/2002  
TIME: 18:59:17

Input Set : A:\USSEQLIST2.txt  
Output Set: N:\CRF3\02132002\I889344A.raw

233	tgtacctcag accatatgag cactatcgaa gaacgca	37
235	<210> SEQ ID NO: 11	
236	<211> LENGTH: 78	
237	<212> TYPE: DNA	
238	<213> ORGANISM: Unknown	
240	<220> FEATURE:	
241	<223> OTHER INFORMATION: Oligonucleotide designed to introduce Q tag	
243	<400> SEQUENCE: 11	
244	tgatgtcagt caagcttacg gaccggcag aactttggac tgggacagg acagcgcctg	60
245	gtggccgtt atgtaatac	78
247	<210> SEQ ID NO: 12	
248	<211> LENGTH: 12	
249	<212> TYPE: PRT	
250	<213> ORGANISM: Artificial Sequence	
252	<220> FEATURE:	
253	<223> OTHER INFORMATION: Derivative of E. coli ACP protein	
255	<400> SEQUENCE: 12	
256	Leu Ser Leu Ser Gln Ser Lys Val Leu Pro Gly Pro	
257	1 5 10	
259	<210> SEQ ID NO: 13	
260	<211> LENGTH: 92	
261	<212> TYPE: DNA	
262	<213> ORGANISM: Unknown	
264	<220> FEATURE:	
265	<223> OTHER INFORMATION: Oligonucleotide designed to introduce Q tag into	
266	Streptococcus haemophilus FabH gene	
268	<400> SEQUENCE: 13	
269	tatcatatga gcctgtccct gtcccagtcc aaagttctgc cgggtccggg taccctcgag	60
270	ggatccggtt ttgcaaaaat aagttaggtt gc	92
272	<210> SEQ ID NO: 14	
273	<211> LENGTH: 53	
274	<212> TYPE: DNA	
275	<213> ORGANISM: Unknown	
277	<220> FEATURE:	
278	<223> OTHER INFORMATION: Oligonucleotide designed to introduce Q tag into	
279	Streptococcus haemophilus FabH gene	
281	<400> SEQUENCE: 14	
282	ctcagatctg agctcaactag tggatcccta aattgtaga atgagcgtgc ccc	53
284	<210> SEQ ID NO: 15	
285	<211> LENGTH: 364	
286	<212> TYPE: PRT	
287	<213> ORGANISM: Artificial Sequence	
289	<220> FEATURE:	
290	<223> OTHER INFORMATION: Modified sequence of Streptococcus haemophilus FabH	
292	<400> SEQUENCE: 15	
293	Met Gly His His His His His His His His Ser Ser Gly His	
294	1 5 10 15	
295	Ile Glu Gly Arg His Met Ser Leu Ser Leu Ser Gln Ser Lys Val Leu	
296	20 25 30	

## VERIFICATION SUMMARY

PATENT APPLICATION: US/09/889,344A

DATE: 02/13/2002

TIME: 18:59:18

Input Set : A:\USSEQLIST2.txt

Output Set: N:\CRF3\02132002\I889344A.raw

L:40 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:71 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:73 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:75 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:77 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:79 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:81 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:83 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:85 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:87 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:89 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:91 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:93 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:95 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:126 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:128 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:130 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:132 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:134 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:136 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:138 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:140 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:142 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:144 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:146 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:148 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:150 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3